

Jeremy Howard
Las maravillosas y pavorosas consecuencias
de las computadoras que pueden aprender

Antes, si querías que un computador hiciera algo nuevo tenías, normalmente, que programarlo. La programación, para quienes no la han practicado, requiere especificar con el más mínimo detalle cada paso que uno quiere que haga su computador para alcanzar el objetivo. Si quieren hacer algo que no saben hacer por sí mismos entonces están ante un gran reto.

Ese fue el reto al que se enfrentó este hombre, Arthur Samuel. En 1956, quería hacer que su computador pudiera ganarle a las damas. ¿Cómo se puede diseñar un programa, teniendo en cuenta el más mínimo detalle que sea mejor que sí mismo a las damas? Y se le ocurrió una idea: hizo jugar al computador contra sí mismo miles de veces y le hizo aprender a jugar a las damas. De hecho funcionó, ya en 1962, este computador había ganado la competición estatal de Connecticut.

Arthur Samuel fue el padre del aprendizaje automático, y tengo una deuda con él, porque soy un profesional del aprendizaje automático. Fui presidente de Kaggle, una comunidad de unos 200 000 profesionales del aprendizaje automático. Kaggle contribuye con competencias para tratar de resolver problemas anteriores no resueltos, y tuvo éxito cientos de veces. Así que desde esa perspectiva, pude descubrir mucho sobre lo que el aprendizaje automático hizo, puede hacer hoy y lo que podrá hacer en el futuro. Quizás el primer gran éxito del aprendizaje automático en el mercado fue Google. Google demostró que era posible encontrar información usando un algoritmo informático y ese algoritmo se basa en el aprendizaje automático. Desde entonces, ha habido muchos éxitos comerciales del aprendizaje automático. Compañías como Amazon y Netflix usan el aprendizaje automático para sugerir artículos que les puedan interesar comprar, películas que les puedan interesar ver. A veces, es casi escalofriante. Compañías como LinkedIn y Facebook dicen, en ocasiones, cuáles pueden ser sus amigos y uno no tiene ni idea de cómo lo han hecho, y es porque hace uso del poder del aprendizaje automático. Estos son algoritmos que han aprendido como hacerlo a partir de los datos en lugar de ser programados a mano.

Así es también como IBM tuvo éxito en conseguir que Watson ganara dos campeonatos mundiales de "Jeopardy" respondiendo preguntas increíblemente ingeniosas, como esta. [El antiguo "León de Nimrud" se perdió del Museo Nacional de esta ciudad en 2003] También es por eso, que podemos ver los primeros autos sin piloto. Si pueden diferenciar entre, por ejemplo un árbol y un peatón, es algo muy importante. No sabemos diseñar

estos programas manualmente, pero con el aprendizaje automático es posible. De hecho, este auto ha conducido más de un millón y medio de kilómetros sin tener accidentes en carretera.

Así, sabemos que las computadoras pueden aprender y pueden aprender a hacer cosas que de hecho nosotros, a veces, no sabemos hacer, o las pueden hacer mejor que nosotros. Uno de los ejemplos más sorprendentes que he visto en aprendizaje automático ocurrió en un proyecto que dirigía en Kaggle donde un equipo dirigido por un chico llamado Geoffrey Hinton de la Universidad de Toronto ganó un concurso para el descubrimiento automático de medicamentos. Lo extraordinario fue no solo que batiera a todos los algoritmos desarrollados por Merck o la comunidad académica internacional, sino que nadie del equipo tenía experiencia en química o biología o ciencias biológicas y lo hicieron en dos semanas. ¿Cómo lo hicieron? Usaron un algoritmo extraordinario llamado aprendizaje profundo. Fue tan exitoso que tuvo cobertura en The New York Times en un artículo de portada unas semanas después. Este es Geoffrey Hinton a la izquierda. El aprendizaje profundo es un algoritmo inspirado en el cerebro humano y como resultado no tiene limitaciones teóricas en lo que puede hacer. Cuantos más datos y tiempo de cálculo uno le dé mejor funciona.

The New York Times mostró, también en su artículo otro resultado extraordinario del aprendizaje profundo que mostraré ahora. Demuestra que las computadoras pueden escuchar y comprender.

(Vídeo) Richard Rashid: El último paso que quiero dar en este proceso es hablar en chino. La clave es, hemos recopilado una gran información de hablantes de chino y producido un sistema de conversión de texto a voz que toma el texto en chino y lo convierte en lengua oral, luego hemos grabado una hora de mi voz que usamos para modular el texto estándar de conversión de texto a voz para que suene como yo. De nuevo, el resultado no es perfecto. De hecho, hay unos cuantos errores. 結果並不完美 (los resultados no son perfectos) Hay mucho que hacer en esta área. 在這方面有很多工作要做 (hay mucho trabajo que hacer en esta área)

Jeremy Howard: Esto era una conferencia de aprendizaje automático en China. No es usual, en conferencias académicas oír aplausos espontáneos, aunque en las conferencias de TEDx siéntanse libres. Todo lo que han visto es gracias al aprendizaje profundo. Gracias. La transcripción en inglés es aprendizaje profundo. La traducción al chino y el texto arriba a la derecha, es aprendizaje profundo, y la construcción de la voz también es aprendizaje profundo.

Eso es lo extraordinario del aprendizaje profundo. Es un solo algoritmo que parece hacer casi cualquier cosa, y descubrí que un año antes, aprendió a ver. En esta extraña

competición en Alemania llamada Banco de Prueba de Reconocimiento de Señales de Tránsito el aprendizaje profundo ha aprendido a reconocer señales de tránsito como esta. No solo reconoce señales de tránsito mejor que cualquier otro algoritmo, la clasificación mostró que era mejor que las personas, dos veces más bueno que las personas. Para 2011, se da el primer ejemplo de computadoras que pueden ver mejor que las personas. Desde entonces, han ocurrido muchas cosas. En 2012, Google anunció que había hecho que un algoritmo de aprendizaje profundo viera vídeos en YouTube y procesaron la información en 16.000 computadoras al mes y el computador aprendió de manera independiente conceptos como personas y gatos solo viendo los vídeos. Esto se parece mucho al aprendizaje humano. Los humanos no aprendemos porque nos cuenten lo que vemos, sino que aprendemos solos qué son esas cosas. También en 2012, Geoffrey Hinton, que vimos anteriormente, ganó la famosa competición de ImageNet, tratando de averiguar, mirando un millón y medio de imágenes, sobre qué eran estas imágenes. A partir de 2014, tenemos un porcentaje de error por debajo del 6 % en reconocimiento de imágenes. De nuevo, mejor que las personas.

Las máquinas están haciendo un trabajo increíble aquí, y está siendo usadas en la industria. Por ejemplo, Google anunció el año pasado que había cartografiado cada sitio de Francia en dos horas, y lo hizo alimentando con imágenes de las calles, al algoritmo de aprendizaje profundo para reconocer y leer los números. Imaginen lo que se habría tardado antes: docenas de personas, muchos años. Esto también está pasando en China. Baidu es como el Google chino, creo, y lo que ven arriba a la izquierda es un ejemplo de una imagen que subí al sistema de aprendizaje profundo de Baidu, y debajo se puede ver que el sistema ha entendido lo que es esa imagen y encuentra imágenes similares. Las imágenes similares tienen fondos similares de las caras, incluso algunos con la lengua afuera. Esto no es claramente mirar el texto de una página web. Todo lo que descargué eran imágenes. Por lo que ahora tenemos computadoras que entienden lo que ven y por ello pueden buscar bases de datos de cientos de millones de imágenes en tiempo real.

¿Qué significado tiene que las computadoras puedan ver? Bueno, no es solo que las computadoras puedan ver. De hecho, el aprendizaje profundo ha hecho más que eso. Frases complejas y llenas de matices como esta son ahora comprensibles con algoritmos del aprendizaje profundo. Como pueden ver aquí, este sistema basado en el de Stanford que muestra el punto rojo en la parte superior ha comprendido que esta frase expresa sentimientos negativos. El aprendizaje profundo está cercano a la conducta humana al comprender lo que significan las frases y lo que se está diciendo sobre esas cosas. El aprendizaje profundo se ha usado también para leer chino, de nuevo, a un nivel casi de hablante nativo. Este algoritmo, desarrollado en Suiza por gente que no hablaba ni entendía chino. Como dije, usar el aprendizaje profundo es el mejor sistema del mundo para esto, hasta comparándolo con el conocimiento humano.

Este es un sistema que formamos en mi empresa que demuestra todas estas cosas juntas. Estas son imágenes sin texto adjunto, y cuando tecleo aquí frases, entiende, en tiempo real, estas imágenes y comprende de qué se tratan y encuentra imágenes similares al texto que estoy escribiendo. Como pueden ver, entiende mis frases y de hecho entiende estas imágenes. Se que han visto algo como esto en Google, donde puede escribir algo y te lo muestra en imágenes, pero lo que realmente está haciendo es buscar la página web a través del texto. Esto es muy diferente a comprender las imágenes. Esto es algo que las computadoras solo han podido hacer por primera vez hace unos pocos meses.

Así que ahora podemos ver que las computadoras no solo ven sino que pueden leer, y, por supuesto, hemos demostrado que pueden entender lo que oyen. Quizá no sea sorprendente ahora lo que voy a decir, pueden escribir. Aquí hay parte de un texto que generé ayer usando el algoritmo de aprendizaje profundo. Y aquí hay parte de un texto que generó un algoritmo de Stanford. Cada una de estas frases fue generada por un algoritmo de aprendizaje profundo para describir estas imágenes. Este algoritmo nunca había visto a un hombre con camisa negra tocando la guitarra. Ha visto a un hombre antes, ha visto el negro antes ha visto una guitarra antes, pero ha generado de manera independiente esta innovadora descripción de esta imagen. Aquí no estamos ante un comportamiento humano, pero estamos cerca. En las pruebas, las personas prefieren las leyendas generadas por el computador 1 de cada 4 veces. Este sistema tiene ahora solo dos semanas de edad, por lo que posiblemente antes del año que viene, el algoritmo del computador irá más allá del comportamiento humano al paso que van las cosas. Así que las computadoras pueden escribir.

Juntamos todo esto y lleva a oportunidades apasionantes. Por ejemplo, en medicina, un equipo de Boston anunció que habían descubierto decenas de características clínicas relevantes sobre tumores que ayudan a los médicos a hacer un diagnóstico de un cáncer. Algo similar, en Stanford, un grupo anunció que, mirando un tejido con aumento, habían desarrollado una máquina basada en el sistema de aprendizaje que de hecho es mejor que los patólogos humanos prediciendo las tasas de supervivencia de los enfermos de cáncer. En ambos casos, no solo fueron las predicciones más precisas, sino que generaron una nueva ciencia reveladora. En el caso de la radiología, hubo nuevos indicadores clínicos que las personas pueden entender. En este caso de patología, el sistema informático descubrió que las células alrededor del cáncer son tan importantes como las células del cáncer mismo al hacer un diagnóstico. Esto es lo contrario de lo que los patólogos han pensado por décadas. En cada uno de estos casos, fueron sistemas desarrollados por una combinación de expertos médicos y expertos del aprendizaje profundo, pero a partir del año pasado, dimos un paso más allá. Este es un caso de identificación de áreas cancerígenas del tejido humano por microscopio. El sistema que se muestra aquí puede identificar esas áreas de formar más precisa o casi tan precisa como los patólogos

humanos, construido completamente con aprendizaje profundo sin usar experiencia médica por gente que no tenía experiencia en este campo. De manera similar, esta segmentación neuronal. Ahora, podemos segmentar neuronas de forma casi tan precisa como las personas, y este sistema fue desarrollado por aprendizaje profundo usando a gente sin experiencia previa en medicina.

Como yo, alguien sin experiencia previa en medicina, parezco completamente calificado para empezar una empresa médica, y lo hice. Estaba aterrorizado de hacerlo, pero la teoría parecía sugerir que podía ser posible hacer medicina muy útil usando solo estas técnicas de información analítica. Afortunadamente, la recompensa ha sido fantástica, no solo por parte de los medios sino de la comunidad médica, que nos ha apoyado mucho. La teoría es que podemos tomar media parte del proceso médico y convertirlo todo lo posible en análisis de datos, dejando a los médicos en lo que son mejores. Quiero dar un ejemplo. Nos lleva unos 15 minutos crear una nueva prueba de diagnóstico médico y ahora lo demostraré en tiempo real, pero lo he comprimido a 3 minutos cortando algunas partes. En vez de mostrar cómo crear una prueba de diagnóstico médico, mostraré una prueba de diagnóstico de imágenes de autos, porque es algo que todos podemos entender.

Así que, empezamos con un millón y medio de imágenes de autos, y quiero crear algo que pueda dividirlas en el ángulo en el que la foto fue tomada. Estas imágenes están sin etiquetar, así que tengo que empezar desde cero. Con nuestro algoritmo de aprendizaje profundo, se pueden identificar automáticamente áreas de la estructura en estas imágenes. Lo bueno es que la persona y el computador pueden trabajar juntos. Así que la persona, como pueden ver aquí, le está indicando al computador áreas de interés que quiere que el computador pruebe y use para mejorar su algoritmo. Estos sistemas de aprendizaje profundo están en un espacio de 16 000 dimensiones, así, pueden ver aquí cómo el computador rota esto en ese espacio, intentando encontrar nuevas áreas de estructura. Y cuando lo hace con éxito, la persona que lo maneja puede, entonces, señalar las áreas de interés. Aquí, el computador ha encontrado, con éxito, áreas, por ejemplo, ángulos. Conforme avanzamos en este proceso, vamos diciendo, gradualmente, al computador más y más sobre los tipos de estructuras que estamos buscando. Pueden imaginar en una prueba diagnóstica que esto debería ser un patólogo identificando áreas patológicas, por ejemplo, o un radiólogo indicando nódulos potencialmente problemáticos. A veces puede ser difícil para el algoritmo. En este caso, queda algo confuso. Las partes delanteras y traseras de los autos están todas mezcladas. Así que tenemos que ser un poco más cuidadosos, seleccionando manualmente las partes delanteras en contraposición a las traseras, para luego decir al computador que este es una especie de grupo es en el que estamos interesados.

Hacemos esto por un tiempo, nos saltamos un poco, y luego probamos el algoritmo de aprendizaje automático basado en un par de cientos de cosas, y esperamos que haya mejorado mucho. Se puede ver, que han empezado a desvanecerse algunas de estas imágenes, mostrándonos que ya está reconociendo cómo entender por sí mismo algunas de ellas. Entonces podemos usar este concepto en imágenes similares, y usando imágenes similares, como pueden ver, en este punto, el computador puede encontrar solo la parte delantera de los autos. En este punto, la persona puede decir al computador, de acuerdo, sí, has hecho un buen trabajo.

En ocasiones, por supuesto, incluso en este punto sigue siendo difícil separar los grupos. En este caso, incluso después de que dejamos al computador que intente girar esto por un momento, seguimos encontrando que la parte izquierda y derecha de las imágenes están mezcladas. Así que podemos dar, de nuevo, al computador algunas pistas, y decimos, bien, intenta encontrar una proyección que separe los lados izquierdos de los derechos de la manera más precisa usando este algoritmo de aprendizaje profundo. Y dándole esta pista... ah, bien, ha tenido éxito. Consiguió encontrar la manera de pensar estos objetos que está separando estos que están juntos.

Así se entiende la idea aquí. Este es un caso en el que la persona no es reemplazada por un computador, sino que trabajan juntos. Estamos reemplazando algo que solía necesitar de un equipo de 5 o 6 personas durante 7 años por algo que lleva 15 minutos a una sola persona.

Este proceso lleva unas 4 o 5 iteraciones. Ahora pueden ver que tenemos un 62 % de nuestro millón y medio de imágenes clasificadas correctamente. En este punto, podemos empezar, con bastante rapidez, a tomar grandes secciones completas, comprobándolas para asegurarse que no hay errores. Cuando hay errores, podemos hacérselo saber al computador. Usando este tipo de proceso para cada uno de los diferentes grupos, nos colocamos en un índice del 80 % de éxito en la clasificación de un millón y medio de imágenes. En este punto, es solo cuestión de encontrar el pequeño número que no está clasificado correctamente, e intentar comprender el porqué. Usando este enfoque, en 15 minutos alcanza un índice de clasificación del 97 %.

Este tipo de técnica nos permite arreglar un problema mayor, que es que hay una falta de conocimientos médicos en el mundo. El Foro Económico Mundial dice que hay entre 10 y 20 veces de escasez de físicos en el mundo desarrollado, y llevará unos 300 años entrenar a gente suficiente para arreglar el problema. Imaginen que pudiésemos ayudar a aumentar su eficiencia usando estos métodos de aprendizaje profundo.

Estoy muy entusiasmado con las oportunidades. También estoy preocupado por los problemas. El problema aquí es que cada área azul de este mapa es algún sitio donde el

empleo, de servicios es mayor del 80 %. ¿Qué son los servicios? Los servicios son estos. Estas son también las mismas cosas que las computadoras acaban de aprender a hacer. Así que el 80 % del empleo mundial en el mundo desarrollado son cosas que las computadoras acaban de aprender a hacer. ¿Qué significa esto? Bueno, no habrá problema, lo reemplazarán por otros trabajos. Por ejemplo, habrá más trabajos para los científicos de datos. Bueno, realmente no. A los científicos de datos no les lleva mucho tiempo construir estas cosas. Por ejemplo, estos 4 algoritmos fueron creados por el mismo chico. Así que si piensan, todo ha pasado ya antes, hemos visto los resultados en el pasado de cuando surgen cosas nuevas y son reemplazadas por nuevos trabajos. ¿Qué trabajos van a ser? Es muy difícil para nosotros hacer una estimación ya que el comportamiento humano crece a un ritmo gradual, pero ahora tenemos un sistema, aprendizaje profundo, que sabemos que crece en capacidad, exponencialmente. Y aquí estamos. Actualmente, vemos las cosas a nuestro alrededor y decimos: "Las computadoras siguen siendo un poco estúpidas". ¿Verdad? Pero en 5 años, las computadoras estarán fuera de esta gráfica. Así que necesitamos empezar a pensar sobre esta capacidad ahora mismo.

Lo hemos visto anteriormente, por supuesto. En la Revolución Industrial, vimos un cambio en la capacidad gracias a los motores. El asunto es, sin embargo, que tras un tiempo, las cosas se nivelan. Hubo una alteración social, pero una vez que los motores se usaron para generar energía en todas las situaciones, las cosas realmente se establecieron. La Revolución del Aprendizaje Automático va a ser diferente a la Revolución Industrial porque la Revolución del Aprendizaje Automático, nunca se asienta. Cuanto mejores son las computadoras en actividades intelectuales, mejores computadoras se crearán para que mejoren su capacidad intelectual, así que esto va a ser una especie de cambio que nunca antes había experimentado el mundo, por lo que el entendimiento previo de lo posible, es diferente. Esto nos impacta. En los últimos 25 años, la productividad del capital se ha incrementado, la productividad laboral se ha mantenido, incluso ha descendido.

Por lo que quiero que empecemos a discutir esto ahora. Sé que cuando hablo sobre esta situación la gente puede ser despectiva. Bueno, las computadoras no pueden realmente pensar, no tienen sentimientos, no entienden poesía, no entendemos realmente cómo funcionan. Y, ¿qué? Actualmente las computadoras pueden hacer cosas en las que las personas gastan su tiempo y les pagan por ello así pues ahora tenemos que empezar a pensar sobre cómo vamos a ajustar nuestras estructuras sociales y económicas para ser conscientes de esta nueva realidad. Gracias.